Crystal structure of the kinase domain of human vascular endothelial growth factor receptor 2: a key enzyme in angiogenesis

Michele A McTigue*, John A Wickersham, Chris Pinko, Richard E Showalter, Camran V Parast, Anna Tempczyk-Russell, Michael R Gehring, Barbara Mroczkowski, Chen-Chen Kan, J Ernest Villafranca and Krzysztof Appelt

Background: Angiogenesis is involved in tumor growth, macular degeneration, retinopathy and other diseases. Vascular endothelial growth factor (VEGF) stimulates angiogenesis by binding to specific receptors (VEGFRs) on the surface of vascular endothelial cells. VEGFRs are receptor tyrosine kinases that, like the platelet-derived growth factor receptors (PDGFRs), contain a large insert within the kinase domain.

Results: We report here the generation, kinetic characterization, and 2.4 Å crystal structure of the catalytic kinase domain of VEGF receptor 2 (VEGFR2). This protein construct, which lacks 50 central residues of the 68-residue kinase insert domain (KID), has comparable kinase activity to constructs containing the entire KID. The crystal structure, determined in an unliganded phosphorylated state, reveals an overall fold and catalytic residue positions similar to those observed in other tyrosine-kinase structures. The kinase activation loop, autophosphorylated on Y1059 prior to crystallization, is mostly disordered; however, a portion of it occupies a position inhibitory to substrate binding. The ends of the KID form a β-like structure, not observed in other known tyrosine kinase structures, that packs near to the kinase C terminus.

Conclusions: The majority of the VEGFR2 KID residues are not necessary for kinase activity. The unique structure observed for the ends of the KID may also occur in other PDGFR family members and may serve to properly orient the KID for signal transduction. This VEGFR2 kinase structure provides a target for design of selective anti-angiogenic therapeutic agents.

Address: Agouron Pharmaceuticals, 3565 General Atomics Court, San Diego, CA 92121, USA.

*Corresponding author.
E-mail: mctigue@agouron.com

Key words: angiogenesis, kinase insert domain, signal transduction, tyrosine kinase, X-ray structure

Received: **14 October 1998**
Revisions requested: **27 November 1998**
Revisions received: **16 December 1998**
Accepted: **14 January 1999**

Published: **4 March 1999**

Structure March 1999, 7:319–330
http://biomednet.com/elecref/0969212600700319

Introduction

Many physiological events including embryogenesis, organ development, estrus, and wound healing require vascular growth and remodeling [1,2]. In addition to these beneficial processes, angiogenesis is also involved in the proliferation of disease states such as tumor growth, metastasis, psoriasis, rheumatoid arthritis, macular degeneration and retinopathy [3–7]. Of the signaling pathways known to influence vascular formation, that involving vascular endothelial growth factor (VEGF) has been shown to be essential and selective for vascular endothelial cells [8–10]. The therapeutic potential of inhibiting the VEGF pathway has been directly demonstrated by anti-VEGF monoclonal antibodies, which are active against a variety of human tumors [11] and ischemic retinal disease [12].

VEGF is a homodimeric cytokine that is expressed in at least four splice-variant forms of 121–206 residues [10]. Vascular endothelial cells express at least two high-affinity receptors for VEGF: VEGFR1 (also referred to in the literature as Flt-1) and VEGFR2 (also referred to as KDR). Another related receptor, VEGFR3/Flt-4, binds the VEGF-like growth factor, VEGF-C, believed to function in lymphogenesis [13]. These VEGFRs are receptor tyrosine kinases each comprised of an extracellular domain that binds VEGF and contains seven immunoglobulin-like segments, a short membrane-spanning region, and a cytosolic domain possessing tyrosine-kinase activity. The kinase domain directly follows the extracellular and juxtamembrane regions and is itself followed by another domain (referred to here as the post-kinase domain), which may function in binding of other proteins for signal transduction. VEGFR1 and VEGFR2 appear to have different signaling pathways and functions, with VEGFR2 being of primary importance in mitosis of endothelial cells [14–16].

On the basis of sequence homology and overall domain structure, VEGFRs belong to the platelet-derived growth factor receptor family (PDGFR), which also includes PDGFRα, PDGFRβ, the stem cell growth factor receptor

(c-kit), and the colony stimulating factor-1 receptor (CSF-1R; also referred to as c-fms) [17]. Compared to other protein kinases, members of this family contain an insert of approximately 65–97 residues, termed the kinase insert domain (KID), within the kinase domain. Deletion or mutation of the KID from PDGFα, PDGFβ, ckit, and CSF-1R has indicated that this domain is not necessary for intrinsic kinase activity but that it is important for the binding of other proteins involved in signal transduction via autophosphorylation of KID tyrosine residues [18–24]. Although the signaling pathways and the specific role of the KID are still not fully determined for VEGFRs, the VEGFR2 KID does contain two tyrosines that have been shown to be autophosphorylation sites [25].

Since the determination of the first cyclic-AMP-dependent protein kinase (cAPK) structure [26] a variety of protein kinase structures have been reported [27]. Among the receptor protein tyrosine kinases (RPTKs), structures of the kinase domain of the insulin receptor (IRK) [28,29] and the fibroblast growth factor receptor 1 (FGFR1) [30,31] have been determined. We report here the generation, kinetic characterization, and structure determination of a modified kinase domain of VEGFR2 containing 18 residues of the 68-residue KID. This 2.4 Å crystal structure of the phosphorylated VEGFR2 catalytic domain is the first reported structure of a kinase domain of the PDGFR family. This structure provides insights into the orientation of the KID domain of VEGFR2, which may be relevant to other PDGFR family members. Furthermore, as inhibition of VEGFR2 kinase has broad clinical potential, this structure provides a three-dimensional description of the target for structure-based design of small molecule VEGFR2 inhibitors as therapeutic agents.

Results and discussion

Kinetic analysis

The tyrosine kinase domain of human VEGFR2 lacking 50 central residues of the 68 residues of the KID was expressed in a baculovirus/insect cell system. Of the 1356 residues of full-length VEGFR2, this protein construct (VEGFR2Δ50) contains residues 806–939 and 990–1171 of the cytosolic domain (Figure 1). In addition to the 50-residue deletion, VEGFR2Δ50 also contains one point mutation (E990V) within the KID relative to wild-type VEGFR2.

Purified VEGFR2Δ50 was found to efficiently catalyze its autophosphorylation as well as phosphorylation of a poly(E_4Y) exogenous substrate. The kinetic methods and analysis reported here are very similar to those previously reported in detail for another VEGFR2 kinase domain protein construct [32].

Using a spectrophotometric assay kinetic constants were determined for both native, nonphosphorylated and autophosphorylated VEGFR2Δ50 (VEGFR2Δ50P) (Table 1). Kinetic characterization of the native nonphosphorylated protein was possible because, under the low protein concentrations used in the activity assays, no autophosphorylation of the protein was detected. Quantitation of the ATPase activity and the phosphotransferase activity indicated that the two reactions were tightly coupled. No ATPase activity was observed in the absence of a peptide substrate.

Since phosphotyrosines are considered high-energy intermediates, the tyrosine-kinase reaction should, in principle, be reversible. In fact, the enzyme catalyses generation of ATP from ADP and phosphorylated poly(E_4Y) [p-poly(E_4Y)] in the presence of Mg^{2+}. Although capable of catalyzing the reverse reaction, VEGFR2Δ50 clearly exhibits a commitment to forward reaction as evidenced by higher k_{cat} for the forward direction relative to that for the reverse direction: 13.0 and 0.13 s^{-1}, respectively, for VEGFR2Δ50P (Table 1).

It is interesting to note that for both nonphosphorylated and phosphorylated VEGFR2Δ50 the $K_{M(MgADP)}$ is less than $K_{M(MgATP)}$ and $K_{M(p\text{-}polyE4Y)}$ is less than $K_{M(polyE4Y)}$. However, the relative differences in these K_m values are much smaller than the relative increase in k_{cat} values between the forward and reverse reactions such that the k_{cat}/K_M values are larger in the forward direction (Table 1).

Comparison of the kinetic constants for VEGFR2Δ50 with those for a VEGFR2 kinase domain containing the entire KID (cchVEGFR2) [32] shows that the 50-residue KID deletion in VEGFR2Δ50 does not have a significant impact on *in vitro* catalytic activity. For example, k_{cat} for VEGFR2Δ50P and phosphorylated cchVEGFR2 in the forward direction were 13 and 12 s^{-1}, respectively. Similarly, $K_{M(MgATP)}$ for VEGFR2Δ50P and phosphorylated cchVEGFR2 were 150 and 153 mM, respectively.

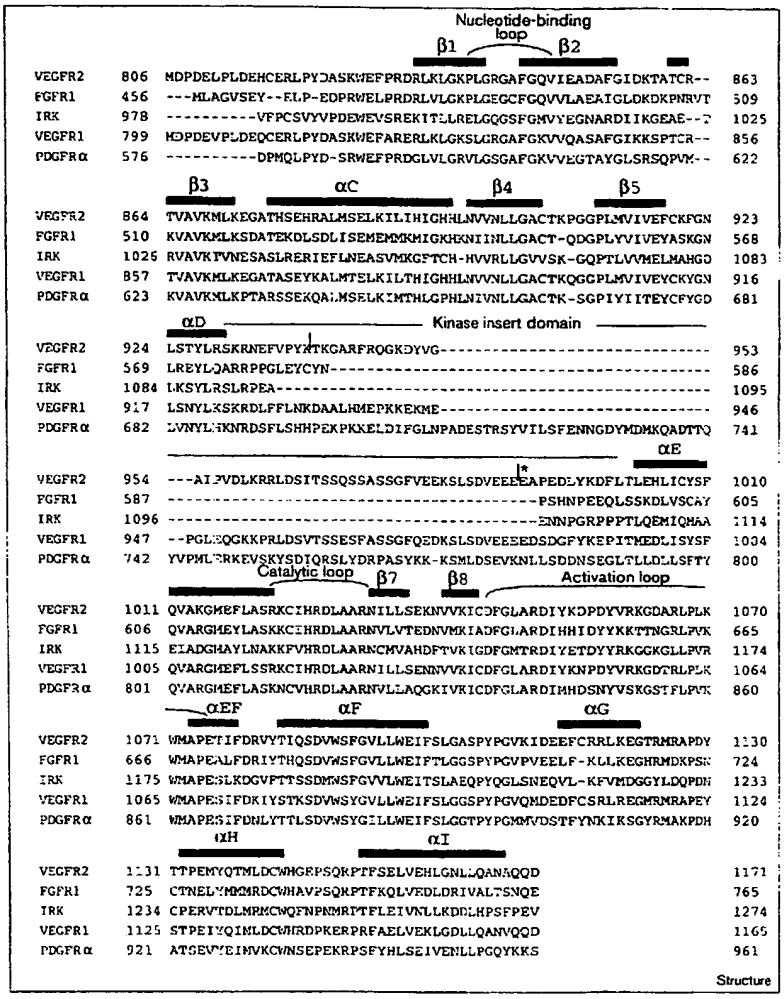
Structure determination

The VEGFR2 KID sequence is hydrophilic and highly charged, containing six lysine, five arginine, eight glutamic acid, and five aspartic acid residues (Figure 1). Initially, several protein constructs containing the VEGFR2 catalytic domain with the entire KID were generated (unpublished results). After exhaustive attempts to crystallize these protein constructs failed to yield even marginal crystals, the VEGFR2Δ50 construct was created to test the idea that the highly charged KID was interfering with crystallization. As determined by dynamic light scattering this VEGFR2Δ50 construct, which eliminated 14 charged residues, exhibited markedly more stability and less multimeric aggregation at various temperatures and protein concentrations than protein constructs containing the entire KID (data not shown).

For crystallization, purified VEGFR2Δ50 was autophosphorylated *in vitro* by incubation with MgATP. Matrix-assisted

Figure 1

Sequence alignment of the kinase domain of VEGFR2 and other RPTKs. Secondary structural elements of VEGFR2Δ50P are indicated: α helices are designated as αC–αI, β strands are designated as β1–β8, and loops are indicated by curves. The site of the 50-residue deletion in VEGFR2Δ50 is indicated by |. The site of the E990V mutation in VEGFR2Δ50 is denoted by an *. Sequences are from VEGFR2 (reported here); FGFR1 (SWISSPROT accession number P11362); IRK (EMBL database number A18657; numbering as in [30]); VEGFR1 (SWISSPROT number P17948); PDGFRα (SWISSPROT number P17948).



laser desorption ionization (MALDI) and nanoelectrospray ionization (NanoESI) mass spectrometry analysis of full-length phosphorylated VEGFR2Δ50 (VEGFR2Δ50P) protein and tryptic-digestion peptides indicates phosphorylation of Y1059. Crystals diffracting to 2.2 Å were obtained of VEGFR2Δ50P in an unliganded state. The crystals belong to the orthorhombic space group $P2_12_12_1$, with one VEGFR2Δ50P molecule in the asymmetric unit.

To check that the protein had not dephosphorylated prior to or during the crystallization process an anti-phosphotyrosine western blot analysis of dissolved crystals was performed (Figure 2). In this analysis samples were run of native nonphosphorylated VEGFR2Δ50, VEGFR2Δ50P taken at two time points during the autophosphorylation reaction, and dissolved crystals. As Figure 2 shows, nonphosphorylated VEGFR2Δ50 did not appreciably bind to the anti-phosphotyrosine antibodies, whereas the autophosphorylated VEGFR2Δ50P and dissolved crystals showed a strong interaction. That the crystals contain phosphorylated protein is also implied by our observation that nonphosphorylated VEGFR2Δ50 did not crystallize under the crystallization conditions reported here for VEGFR2Δ50P.

Initial crystallographic phases were determined by molecular replacement using the structure of the unphosphorylated kinase domain of FGFR1 [30] as a search model. The correctness of the molecular replacement solution was cross-checked by calculating a difference Fourier with a gold cyanide derivative. The derivative data, however, were not included for phase calculations of electron-density maps used to build the model. The structure has

Table 1

Kinetic constants of VEGFR2Δ50

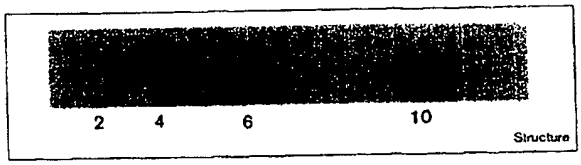
Substrate	Native VEGFR2Δ50			Phosphorylated VEGFR2Δ50		
	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}M^{-1}$)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}M^{-1}$)
Forward reaction						
MgATP	0.597	5.5	9.2×10^3	0.153	13.3	87×10^3
poly(E_4Y)	9.4		5.8×10^2	2.1		63×10^2
Mg^{2+}	31.0		1.8×10^2	6.8		20×10^2
Reverse reaction						
MgADP	0.127	0.1	7.9×10^2	0.056	0.13	23×10^2
p-poly(E_4Y)	1.3		7.7×10^1	1.0		13×10^1

been refined to an R factor of 20.2% for 8–2.4 Å data ($F_o > 2\sigma$). VEGFR2Δ50P residues for which backbone atoms were not modeled due to disorder include the N-terminal residues 806–819, residues 990–997 of the KID, residues 1048–1063 of the activation loop, and residues 1169–1171 of the C terminus. Structure-determination statistics are included in Table 2.

Overall kinase fold

Analogous to previously reported structures of both serine/threonine and tyrosine protein kinases, VEGFR2Δ50P is folded into two lobes, with catalysis of phosphotransfer taking place in a cleft between the two lobes (reviewed in [27,33]). A Cα trace of the VEGFR2Δ50P structure is shown in Figure 3a. Kinase secondary structural elements are designated (Figure 1) according to the convention originally given for cAPK [26]. The N-terminal lobe (approximately residues 820–920) folds into a twisted β sheet with one α helix (αC). The β structure comprises five antiparallel strands (β1–β5), three of which (β1–β3) are highly curved and curl over the other two strands (β4–β5). The larger C-terminal domain (approximately residues 921–1168) contains two antiparallel β strands (β7–β8), which lie at the top of the C-terminal domain adjacent to the N-terminal β sheet. Seven α helices (αD, αE, αE-F, αF, αG, αH, αI) form the remaining core of the C-terminal domain. Like other kinases, VEGFR2Δ50P also contains functionally important loop regions: the glycine-rich nucleotide binding loop (residues 841–846), the catalytic loop (residues 1026–1033) and the activation loop (residues 1046–1075; Figures 1 and 3a).

Figure 2



Phosphorylation state of VEGFR2Δ50 crystals. A cross section of an antiphosphotyrosine western blot is shown. Regions of the blot not shown did not contain visible bands. Lane 2, purified VEGFR2Δ50; Lane 4, purified VEGFR2Δ50 after 20 min of autophosphorylation; Lane 6, purified VEGFR2Δ50 after 40 min of autophosphorylation; Lane 10, dissolved crystals.

Of the reported kinase structures, the VEGFR2Δ50P structure resembles most closely that of the catalytic domain of FGFR1 [30] with which it shares approximately 55% sequence identity (Figure 1). As the two molecules in the crystallographic asymmetric unit of the FGFR1 structure solution are very similar, comparisons to VEGFR2Δ50P will primarily be described only for the FGFR1 molecule A. Least-squares superposition of 82 Cα positions of strands (β1–β5) of the N-terminal lobe or 152 Cα positions of helices (αD, αE, αF, αG, αH, αI) of the C-terminal lobe between FGFR1 and VEGFR2Δ50P result in respective root mean square deviation (rmsd) values of 0.40 Å and 0.52 Å. A relative rotation of approximately 5° between the two lobes results in the interlobe cleft of VEGFR2Δ50P being slightly larger and more open. Measurement of distances between equivalent Cα atoms (K523 and R675 of FGFR1, S877 and R1080 of VEGFR2Δ50P) at the ends of the cleft reveal that this distance is 25.3 Å in VEGFR2Δ50P in comparison with 23.2 Å in FGFR1. This is, however, a minor difference in comparison with much larger relative lobe rotations observed among kinase structures in various complexes and phosphorylation states [27]. For example, the interlobe orientation seen here for VEGFR2Δ50P is in an approximately 20° more open conformation than that seen in the ternary complex structure of the phosphorylated kinase domain of IRK bound to the ATP analog, adenylyl imidodiphosphate (AMP-PNP) and a peptide substrate [29] (Figure 3c).

Although the β-strand positions of the N-terminal lobe agree well between VEGFR2Δ50P and FGFR1, the structures do diverge significantly at the N-terminal residues preceeding the first conserved region, which starts at residue W827 (Figure 3a and 3b). The first 14 residues (M806–E819) of VEGFR2Δ50P are completely disordered and the next seven residues (L820–R826) form an extended-loop structure. It is likely that residues 806–819 do not form part of the active kinase region but are instead part of, or

Table 2

VEGFR2Δ50P structure determination statistics

Data set	Native (3)	Native (1)	Native (2)	$KAu(CN)_2$
Data resolution (Å)	15–2.2	20–2.9	15–2.4	15–3.1
R_{sym}*† (%)	5.2 (19.6)	8.4 (26.6)	7.0 (21.9)	7.1 (19.5)
Completeness (%)	93.0 (81.0)	97.5 (98.4)	98.8 (98.8)	96.5 (95.0)
Temperature (°C)	−186	room (~21)	4	4
Unit cell a (Å)	95.41	97.10	98.52	97.71
Unit cell b (Å)	96.04	96.94	96.50	96.97
Unit cell c (Å)	38.22	38.63	38.56	38.52
Refinement resolution (Å)	8–2.4	–	–	–
Refined R‡§ (%)	20.2	–	–	–
Rfree¶ (%)	27.3			

*$R_{sym} = 100 \times \Sigma_{hkl}\Sigma_i|I_i\,(hkl) - \langle I(hkl)\rangle| \,/\, \Sigma_{hkl}\Sigma_i I_i\,(hkl)$. †Value in parentheses is for highest resolution shell: Native (3), 2.28–2.20; Native (1), 3.02–2.90; Native (2), 2.49–2.40; $KAu(CN)_2$, 3.21–3.10. ‡$R = 100 \times \Sigma_{hkl}||F_o(hkl)| - |F_c\,(hkl)|| \,/\, \Sigma_{hkl}|F_o\,(hkl)|$ where F_o and F_c are the observed and calculated structure factors, respectively ($F_o > 2\sigma$). §Model includes 275 protein residues and 182 water molecules. ¶Calculated using a test set of 7% of the Native (3) data and subjecting the final refined model to simulated annealing and positional refinement with this test set not included in the minimization calculations.

are adjacent to, the juxtamembrane region of VEGFR2. Residues 820–826 do seem to be part of the kinase domain as analogous residues are also ordered in the structures of FGFR1, IRK, and the non-receptor tyrosine kinase Lck [34]. Other differences between the VEGFR2Δ50P structure and other kinase structures occur at the KID and the activation loop (discussed below).

Catalytic loop and ATP-binding site

In protein kinases, the loop between αE and β7 has been termed the catalytic loop as it contains an invariant aspartic acid residue (D1028) that is essential for catalysis of the phosphotransfer reaction [27]. This aspartic acid is part of a stretch of residues (H1026–N1033) whose sequence HRD-LAARN (single-letter amino acid code) is highly conserved among protein tyrosine kinases. In VEGFR2Δ50P the backbone position and most sidechain positions of this loop are similar to those in the unliganded FGFR1 and ternary phosphorylated IRK (IRKP) complex structures. As seen in these previous structures, the sidechain carboxylate of the catalytic loop aspartic acid (D1028) is hydrogen-bonded to the sidechains of the conserved arginine (R1032) and asparagine (N1033; Figure 4).

The ATP-binding site of protein kinases lies at the cleft between the N- and C-terminal lobes (Figure 3c). For

Figure 3

Overall fold of VEGFR2Δ50P, FGFR1, and IRKP. Backbone representation of structures of the kinase domains of **(a)** VEGFR2 (VEGFR2Δ50P), **(b)** FGFR1 ([30], molecule A of Protein Data Bank [PDB] entry 1FGK), and **(c)** IRKP ([29], PDB entry 1IR3). The views shown in (a), (b), and (c) are identical views generated from superpositions of the C-terminal domains. The positions of the termini are denoted by N and C. In (a) β strands are designated as β and α helices are indicated by α. The nucleotide-binding loop (orange), KID (purple), and activation loop (yellow) are highlighted. In (c) the bound AMP-PNP is shown in green and the peptide substrate is shown in red. The figure was prepared with the program INSIGHT II [53].

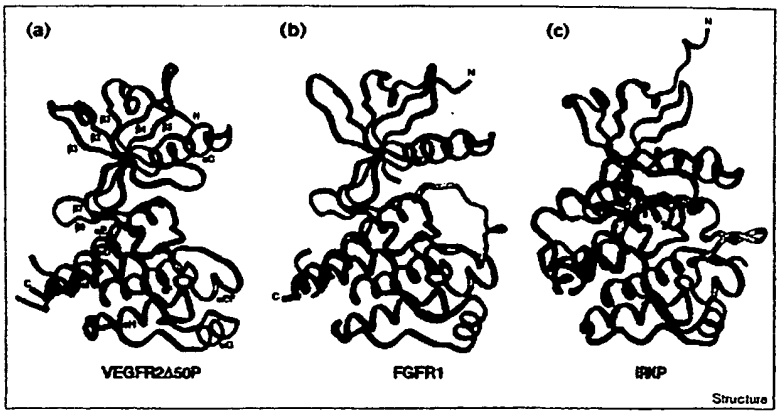
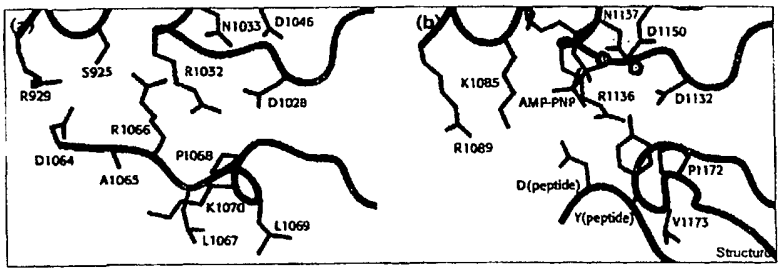


Figure 4



Catalytic site of VEGFR2Δ50P and IRKP. Cross section of the catalytic site of **(a)** VEGFR2Δ50P and **(b)** IRKP [29] structures. Protein Cα ribbon representations are shown in light blue. In (b) the Cα ribbon representation of the bound peptide substrate is shown in orange. Specific atoms are colored by element type: carbon (green), oxygen (red), nitrogen (blue), phosphorous (purple) and magnesium ions (yellow). (a) includes only protein atoms. (b) includes protein atoms, peptide atoms, AMP-PNP atoms, and Mg^{2+} ions. The figure was generated using INSIGHT II [53].

VEGFR2Δ50P, the residues forming this site consist primarily of residues E917–N923 that join the two lobes, and residues L840–I849 that include parts of β1, β2 and the glycine-rich loop of G841–G846. The glycine-rich loop, also referred to as the nucleotide binding loop, is a flexible segment whose position differs among kinase structures in various activated and liganded states. In VEGFR2Δ50P this loop is fairly well ordered and all atoms could be modeled, with the exception of the sidechains of R842 and F845. The relative position and conformation of this loop is similar to that observed in the unliganded FGFR1 structure. However, this position is markedly different from that in the IRKP ternary complex structure in which the approximately 20° relative rotation of the N- and C-terminal lobes results in the glycine-rich loop being 5 Å closer to the C-terminal lobe than in VEGFR2Δ50P.

In reported kinase structures with bound ATP or an ATP analog, the adenine ring makes two conserved hydrogen bonds with the protein backbone. In the structure of FGFR1 with adenylyl methylenediphosphate (AMP-PCP) bound [30], these hydrogen bonds are between the adenine NH_2 and the backbone C=O of E562 (E917 VEGFR2Δ50P) and between the adenine N1 and the backbone NH of A546 (C919 VEGFR2Δ50P). Although the structure presented here does not contain a bound nucleotide, the similarities in the positions of these backbone atoms to those in FGFR1 indicate that these hydrogen bonds would be formed in a VEGFR2Δ50P–ATP complex and therefore the adenine is expected to bind in a similar position (Figure 5).

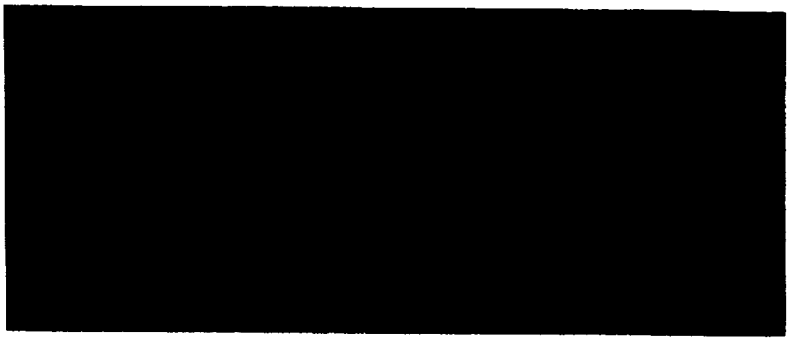
Variation in the ATP-binding sites of kinases involved in disease is of considerable importance in the design of selective ATP-competitive inhibitors as therapeutics. A comparison of the ATP-binding sites of FGFR1 and VEGFR2Δ50P reveals that while the overall architecture of the site is conserved, several sequence differences result in differences in the shape of the accessible area for ligand binding. Specific sequence differences between FGFR1 and VEGFR2 in this site include V899 (I545 FGFR1), F918 (Y563 FGFR1), C919 (A564 FGFR1), and C1045 (A640 FGFR1; Figure 5). Similarly, comparison to the ternary IRKP complex structure reveals variation in the adenine site at V916 (M1076 IRK), F918 (L1078), C919 (M1079 IRK), L1035 (M1139 IRK), and C1045 (G1149 IRK). Even greater sequence and structural variation in the adenine site is seen when the VEGFR2Δ50P structure is compared to serine/threonine-kinase structures, suggesting

Figure 5

Nucleotide-binding site of VEGFR2Δ50P and FGFR1. Stereoview showing the Cα trace and some sidechains of a superposition of the nucleotide-binding sites of the VEGFR2Δ50P and the FGFR1–AMP-PCP complex ([30]; molecule B) structures. The superposition was performed using Cα positions of helices (D, E, F, G, H, and I) of the C-terminal lobes. Carbon atoms of VEGFR2Δ50P are shown in yellow and carbon atoms of FGFR1 are shown in purple. The coloring for other protein atoms is: oxygen (red), nitrogen (blue), and sulfur (green). The AMP-PCP in the FGFR1 structure is depicted in orange. Labels correspond to VEGFR2Δ50P residues. The figure was created with the program Xfit [50].

Figure 6

Electron-density map of the kinase insert domain area of VEGFR2Δ50P. Stereoview of a $2F_o-F_c$ map computed at 2.4 Å and contoured at 1.2σ, and superimposed with the refined model. Carbon atoms are yellow, oxygen atoms red, and nitrogen atoms are blue. Water molecules are depicted as red crosses. The figure was created with Xfit [50].



that these differences could be exploited in the design of selective ATP-competitive inhibitors.

Activation loop

Protein kinases contain a large flexible loop, referred to as the activation loop (A loop), whose conformation is postulated to regulate kinase activity (Figure 3). In many kinases the conformation of the A loop is controlled by the phosphorylation of specific A-loop residues [27]. The loop can be generally defined as beginning with the conserved residues DFG and ending at the conserved APE sequence [27]. In VEGFR2 this segment corresponds to D1046–E1075 and contains two tyrosines (Y1054 and Y1059). Both Y1054 and Y1059 were found to be autophosphorylation sites when the cytosolic domain of VEGFR2 was expressed in *Escherichia coli* [25]. Using the *in vitro* autophosphorylation protocol described here for VEGFR2Δ50, a stable phosphorylation site is indicated at Y1059; however, no evidence of phosphorylation of Y1054 was detected.

In the VEGFR2Δ50P structure presented here, the A loop appears quite mobile and interpretable electron density was not present for most of the central portion of the loop (G1048–G1063). This disorder is consistent with mobility of the A loop deduced from other kinase structures. For example, of the two molecules in the asymmetric unit of the unphosphorylated FGFR1 kinase structure, the center of the A loop has relatively high temperature factors in molecule A and is completely disordered in molecule B. Although residues 1048–1063 could not be modeled in VEGFR2Δ50P, unambiguous electron density was present for residues D1064–E1075, clearly indicating that these residues adopt a conformation similar to that observed in the unphosphorylated FGFR1 structure. The segment of D1064–P1068 has an extended structure that lies adjacent to the catalytic residues D1028 and R1032 (Figure 4a). Comparison with the structure of the MgAMP-PNP–peptide–IRKP complex indicates that the position of R1066–P1068 in this VEGFR2Δ50P structure is inhibitory to substrate binding. P1066 occupies the equivalent space allocated to the tyrosine sidechain of the peptide substrate in the ternary IRK3P complex structure. The conformation of residues L1069–E1075 is similar to that in the ternary IRKP complex structure. However, there is a complete directional change at P1068 (P1172 IRK) between the two structures. In the IRK structure, residues N-terminal to this proline are directed toward αEF, whereas in VEGFR2Δ50P they are directed toward αD on the opposite side of the protein (Figures 3 and 4).

It is somewhat surprising that despite the phosphorylation of Y1059 prior to crystallization, the conformation seen here for residues D1064–P1068 is similar to the inhibitory conformation observed for analogous residues in the unphosphorylated FGFR1 structure. Y1059 in VEGFR2Δ50 corresponds to a relatively conserved phosphorylation site among protein tyrosine kinases. In the ternary IRKP complex structure and the phosphorylated Lck structure [34], the tyrosine at this position (Y1163 IRK, Y394 Lck) is phosphorylated and the A loop has a non-inhibitory conformation similar to that observed in a phosphorylated cAPK ternary complex structure [35]. The interactions the phosphate group at this position makes with other protein residues are believed to help stabilize an A-loop conformation that allows substrate and ATP binding [27,29]. However, as the VEGFR2Δ50P structure described here does not exhibit a similar open A-loop conformation but rather has an inhibitory conformation with much of the loop disordered, it is possible that the monophosphorylated A loop of VEGFR2Δ50P exists in a dynamic equilibrium involving several conformations and that the conformation observed here is the one most favored in this crystal environment.

Kinase insert domain

The kinase insert domain occurs in the kinase C-terminal lobe and connects helices αD and αE. In VEGFR2 this region corresponds to a 68-residue stretch from N933 to L1000 (Figure 1). The lack of effect on intrinsic kinase activity (noted above) of deletion of residues T940–E989

is perhaps not surprising, as the ends of the KID domain are relatively far away (approximately 35–40 Å) from the catalytic site and are on the opposite side of the protein from the position of the activation loop (Figure 3). These results agree with those for the CSF-1R kinase in which deletion of 58 of the 64 residues of the CSF-1R KID only decreased its ability to phosphorylate a peptide substrate by 10% [18]. Deletion of the entire 98 KID residues of βPDGFR, however, resulted in an 80% decrease in kinase activity towards a peptide substrate [36]. Considered together, these results suggest that the majority of the KID is not required for catalysis but that a small number of residues must be present to form a linker between αD and αE to maintain a competent kinase structure.

In the VEGFR2Δ50P structure, following αD, residues N933–P937 form a loose turn and an extended strand whose ends are roughly perpendicular to the axes of αD and αI at the C terminus. In difference Fourier maps the electron density is strong and clear for residues N933–P937 and becomes weak for Y938 and K939 (sidechains of Y938 and K939 are not modeled; Figure 6). The 50-residue deletion in VEGFR2Δ50 directly follows K939 so that the residue immediately C-terminal to K939 is V990, maintaining the residue numbering in full-length VEGFR2. Residues V990–K997 are disordered and interpretable electron density begins again at D998. Residues D998–T1001 then form a short strand that joins αE at residue L1002 (Figures 6 and 7).

An interesting feature of these two strands at the N- and C-terminal ends of the KID is that they form a pseudo two-stranded parallel β-sheet structure that is quite different from the conformations seen in this region of other kinase structures. The two ends of the KID thus make a variety of interactions that may help to stabilize the overall conformation and position of this domain in VEGFR2. The sidechain of K931 makes an ionic interaction with the sidechain of E934 and also makes a hydrogen bond with the backbone carbonyl of D998 (Figure 7). Hydrogen-bonding interactions between the strands include E934 backbone C=O to L1000 NH, V936 NH to L1000 C=O, and P937 C=O to L1002 NH. In addition to these polar interactions, the sidechains of F935, P937, and L1000 are involved in extensive hydrophobic contacts. The sidechain of F935 is nestled in a hydrophobic pocket formed by the sidechains of L928, P937, L1000, L1002, L1005, L1101, and Y1130 (Figures 6 and 7). The L1000 sidechain also packs against the sidechains of Y927, K931, H1004, and Y1008.

The length, composition, and structure of the connecting region between αD and αE varies considerably among kinases. For brevity, only comparisons to the IRK and FGFR1 receptor tyrosine kinase structures will be described here. In IRK this connection stretches twelve residues, of which five are proline, from P1093 to P1104. This proline-rich segment is ordered and exhibits similar loop-like conformations in the unphosphorylated apoenzyme structure [28] and in the phosphorylated ternary complex structure [29]. The most significant divergence in the KID position in IRK and VEGFR2Δ50P begins immediately following αD at N933 (P1093 IRK). In IRK, residues P1093–G1100 form a loop that folds towards the N-terminal domain. By contrast, the path taken by N933–K939 in VEGFR2Δ50P is towards the C-terminal helix αI (Figure 3a).

In FGFR1 the connection between αD and αE is 18 residues long (P578–L595) and shares little sequence homology with VEGFR2 or IRK. In FGFR1 this region is apparently quite mobile because, for the two molecules in the crystallographic asymmetric unit of the FGFR1 structure, residues 580–591 are disordered in molecule A and residues 578–593 are disordered in molecule B [30]. Although most of this segment is disordered in FGFR1, the ordered ends follow a path more similar to IRK than to VEGFR2Δ50P.

Sequence comparisons with other members of the PDGFR family suggest that some of these kinases may have conformations for the beginning and end residues of the KID

Figure 7

Stereo cross section showing the ordered residues of the KID of VEGFR2Δ50P. Carbon atoms are yellow, oxygen atoms are red, nitrogen atoms are blue, and sulfur atoms are green. The view is rotated roughly 180° from that of Figure 6. The figure was created with Xfit [50].

similar to those seen here for VEGFR2Δ50P. In particular, there is significant sequence homology in the first four KID residues (N933–V936) that make the turn towards the C terminus. Of these residues, a phenylanine at the position of F935, which makes extensive hydrophobic interactions, may be particularly important for structural integrity as a phenylalanine occurs at this position in sequence alignments with VEGFR1, VEGFR3, αPDGFR, βPDGFR, and ckit [37–40].

Biological implications

The crystal structure and kinetic characterization of a protein construct (VEGFR2Δ50) of vascular endothelial growth factor receptor 2 (VEGFR2), containing the kinase core and 18 of the 68 residues of the kinase insert domain (KID), provide several biological insights. The high similarity in overall fold and catalytic-residue positions to the fibroblast growth factor receptor 1 (FGFR1) and insulin receptor kinase (IRK) kinase domain structures suggests that the catalytic mechanism of VEGFR2, and other platelet-derived growth factor receptor PDGFR-family members, is consistent with other tyrosine kinases. Tyr1059, of the VEGFR2 kinase activation loop, was confirmed to be an autophosphorylation site. The activation loop is predicted to exist in a dynamic equilibrium of several conformations, one of which is similar to a conformation observed in FGFR1 and is inhibitory to substrate binding.

As deletion of the 50 central residues of the KID did not affect autophosphorylation or kinase activity on an exogenous substrate, this domain does not seem to have a direct function in the phosphotransfer reaction. Although not directly involved in catalysis, evidence suggests that the KID functions in signal transduction as a binding site for cellular proteins. The VEGFR2 KID contains three tyrosines (Y938, Y951, and Y996) of which Y951 and Y996 have previously been shown to be autophosphorylation sites. The determination of these autophosphorylation sites as binding sites for other proteins has yet to be reported for VEGFR2; however, phosphatidyl inositol 3 kinase, Grb2, and SHP-2 have been to shown to bind to phosphotyrosines in the KID of other PDGFR family members [41].

In the structure reported here, the residues forming the two ends of the VEGFR2 KID interact with each other to form a pseudo-β substructure not observed in previous protein kinase structures. This substructure is stabilized by a large number of interactions and is expected to also occur when the entire KID is present. The close proximity of this pseudo-β substructure to the post-kinase domain implies that the KID and post-kinase domains pack near to each other in full-length VEGFR2 and perhaps in other PDGFR family members. In this case, the pseudo-β structural element may serve to orient the KID to pack near to the post-kinase domain. As both the KID and the post-kinase domains of these receptor protein tyrosine kinases (RPTKs) are believed to be involved in the binding of other proteins, it is possible that having the KID and post-kinase domains in near proximity may facilitate formation of multimeric complexes for signal transduction.

Materials and methods

Cloning

The coding sequence [42] for the cytoplasmic domain of the VEGFR2 was amplified by the polymerase chain reaction (PCR) [43] from a human aorta cDNA pool (Clontech, Palo Alto, CA). Two overlapping sequences were amplified independently: Vcyt (residues M806–V1356), which represented the entire cytoplasmic domain, and Vcat (residues C817M–G1191), with boundaries based upon a primary amino acid sequence alignment with the insulin receptor kinase catalytic domain [44]. The PCR oligonucleotide primer sequences for Vcyt were; Vcyt5: 5′-CAGCATATGGATCCAGATGAACTCCCATTGG-3′ and Vcyt: 5′-GCGGTCGACTTAAACAGGAGGAGAGCTCAGTGTG-3′. The PCR oligonucleotide primer sequences for the Vcat were; Vcat5 5′-GCACATATGGAACGACTGCCTTATGATGCCAGC-3′ and Vcat3: 5′-CCTGTCGACTTATCCAGAATCCTCTTCCATGCTCAAAG-3′. The amplified DNA was digested with the restriction enzymes *Nde*I and *Sal*I, ligated into the *E. coli* plasmid pET24a (Novagen Madison, WI), and the sequence verified. When compared to the original sequence in GENBANK, accession number 346345, two nucleotide differences were noted that resulted in amino acid changes (Glu848→Val and Asn835→Lys) in both Vcyt and Vcat. Our sequence agrees with subsequent VEGFR2 GENBANK submissions (accession numbers 2655412 and 3132833).

Mutations were introduced by oligonucleotide site-directed mutagenesis [45] using the Muta-Gene in vitro Mutagenesis Kit from Bio-Rad (Hercules, CA). The Vcat DNA fragment was subcloned from the pET24a vector using an *Nde*I–*Xho*I digest into the vector pMGH4 [46,47] and this vector was used to generate the single-stranded DNA (ssDNA) uracil-containing template (minus strand) in *E. coli* strain CJ236 supplied in the kit. An oligonucleotide (5′-GCTCAGCAGGATTGATAAGACTACATTGTTC-3′) was designed to create a construct, Vcat(ΔG1172–G1191), which truncated the C terminus to residue D1171. Another oligonucleotide (5′-GAATTTGTCCCCTACAAGGAAGCTCCTGAAGATCTG-3′) was designed to delete the central 50 residues (residues T940–E989) of the KID, on the basis of a sequence alignment with FGFR1 [30]. Sequence analysis detected an inadvertent Glu990→Val mutation. All DNA modification and restriction enzymes were purchased from New England Biolabs and oligonucleotides were purchased from Genosys Biotechnology.

The DNA encoding the VEGFR2Δ50 protein construct was made in several steps to combine the necessary mutations into the baculovirus expression vector pAcSG2 (Pharmingen, San Diego, CA). Step 1: the coding region for Vcyt was subcloned from the pET24a vector into the *Nco*I–*Kpn*I sites of vector pAcSG2. Step 2: a 2358 base pair (bp) *Sca*I–*Bgl*II DNA fragment from plasmid pMGH4–Vcat(ΔT940–E989, E990V) was ligated to a 1695 bp *Bgl*II–*Sca*I DNA fragment from pMGH4–Vcat(ΔG1172–G1191) creating a pMGH4–Vcat(ΔT940–E989, E990V, ΔG1172–G1191) vector. Step 3: a 913 bp *Bst*EII–*Eag*I DNA fragment, a pMGH4–Vcat(ΔT940–E989, E990V, ΔG1172–G1191), was ligated to a 3290bp *Eag*I–*Bst*EII DNA fragment from pAcSG2–Vcyt, creating pAcSG2–Vcyt(ΔT940–E989, E990V, ΔG1172–G1191). This final DNA construct was sequenced, verified through the entire coding region and it was confirmed to contain only these known mutations from the wild-type sequence (Figure 1).

The pAcSG2–Vcyt(ΔT940–E989, E990V, ΔG1172–G1191) DNA, expressing the VEGFR2Δ50 protein, was transfected into *Spodoptera frugiperda* (Sf9) cells with linearized baculovirus DNA according to the

protocol of the manufacturer (Pharmingen, San Diego, CA). Single plaques were isolated from this transfection and high-titer stocks generated. All stocks were examined by isolation of baculoviral DNA and PCR amplification of the insert using the polyhedrin forward and reverse primers (Invitrogen). Sf21 cells were infected at 1–1.5 million cells/ml at a multiplicity of infection of 5 for 72 h and harvested by centrifugation.

Protein purification

Cell pellets were lysed by dounce homogenization and sonication in 20 mM Tris pH 8.0, 20 mM NaCl, 5 mM dithiothreitol (DTT), and 5% (v/v) glycerol. The lysate was centrifuged for 50 minutes at 35,000 rpm in a Ti45 rotor. The soluble fraction was loaded onto a 40 ml Q-30 anion-exchange column (Pharmacia, Piscataway, NJ) and eluted with a 20–600 mM NaCl gradient in 20 mM Tris pH 8.0, 5 mM DTT, and 5% (v/v) glycerol over 20 column volumes. VEGFR2Δ50 was pooled by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis and by the presence of kinase activity as measured against gastrin peptide substrate (Boehringer Mannheim, Indianapolis, IN). Pooled material was loaded onto a 40 ml hydroxyapatite (Bio-Rad, Hercules, CA) column and washed extensively with 20 mM Tris pH 8.0, 50 mM NaCl, 5 mM DTT, and 5% glycerol. Protein was eluted using a 500 ml linear gradient from 0–50 mM potassium phosphate pH 8.0, 50 mM NaCl, 5 mM DTT, and 5% glycerol. VEGFR2Δ50 was pooled by SDS–PAGE analysis and by the presence of kinase activity as measured against the gastrin peptide. Material from this column was then diluted 1:1 with 20 mM Tris pH 8.0, 20 mM NaCl, 5 mM DTT, and 5% glycerol and loaded onto an 8 ml Q-15 anion-exchange column (Pharmacia). Protein was eluted using a 180 ml linear NaCl gradient (20–175 mM) in 20 mM Tris pH 8.0, 5 mM DTT, and 5% glycerol. VEGFR2Δ50 was pooled as described above. $(NH_4)_2SO_4$ (4M) was added to the pool to a final concentration of 0.6 M and the pooled material was loaded onto a 10 ml HP-phenyl sepharose column (Pharmacia). VEGFR2Δ50 was eluted using a 200 ml linear reverse gradient from 0.6–0 M $(NH_4)_2SO_4$ in 20 mM Tris and 5 mM DTT. Purified VEGFR2Δ50 was buffer exchanged into 50 mM Hepes pH 7.5, 10 mM DTT, 10% glycerol, and 25 mM NaCl over a 500 ml G-25 column (Pharmacia) and concentrated to 1 mg protein/ml through a 10 kDa cutoff polysulfone membrane (Millipore, Bedford, MA). Final material was aliquoted and flash frozen in liquid N_2 and stored at −70°C.

Kinetic assays

Kinetic assays of phosphorylated VEGFR2 were performed with purified VEGFR2Δ50 that was autophosphorylated under the following conditions: protein (4 mM), ATP (3 mM), $MgCl_2$ (40 mM), DTT (5 mM), in Hepes (100 mM), 10% glycerol, pH 7.5, at 4°C for 1 h. Analysis of data was carried out as described in [32].

Coupled spectrophotometric assay for the forward reaction

Tyrosine-kinase assays were monitored using a Beckman DU 650 spectrophotometer. Production of ADP was coupled to oxidation of NADH using phosphoenolpyruvate (PEP) through the actions of pyruvate kinase (PK) and lactic dehydrogenase (LDH). The oxidation of NADH was monitored by following the decrease in absorbance at 340 nm ($\varepsilon 340 = 6.22\ cm^{-1}mM^{-1}$). Typical reaction solutions contained 1 mM PEP, 250 mM NADH, 50 units of LDH/ml, 20 units of PK/ml, 5 mM DTT, in 200 mM Hepes, pH 7.5, and indicated concentrations of poly(E_4Y_1) (Sigma, St. Louis, MO), ATP and $MgCl_2$. Assays were initiated with either VEGFR2Δ50P (40 nM) or VEGFR2Δ50 (70 nM). For nonphosphorylated enzyme, $K_{M(MgATP)}$ was determined by varying the ATP concentration at fixed concentrations of Mg^{2+} and poly(E_4Y) (60 and 20 mM, respectively). For phosphorylated enzyme, $K_{M(MgATP)}$ was determined by varying the ATP concentration at fixed concentrations of Mg^{2+} and poly(E_4Y) (25 and 5 mM, respectively). $K_{M(polyE4Y)}$ for the nonphosphorylated enzyme was determined by varying the concentration of poly(E_4Y) at fixed concentrations of Mg^{2+} and ATP (65 and 2 mM, respectively). $K_{M(polyE4Y)}$ for the phosphorylated enzyme was determined by varying the concentration of poly(E_4Y) at fixed concentrations of Mg^{2+} and ATP (25 and 5 mM, respectively). For nonphosphorylated enzyme, $K_{M(Mg2+)}$ was determined by varying the Mg^{2+} concentration at fixed concentrations of ATP and poly(E_4Y) (2 and 18 mM, respectively). For phosphorylated enzyme, $K_{M(Mg2+)}$ was determined by varying the Mg^{2+} concentration at fixed concentrations of ATP and poly(E_4Y) (1 and 5 mM, respectively).

Coupled spectrophotometric assay for the reverse reaction

ATP generation was coupled to production of NADH via the action of hexokinase (HK) and glucose-6-phosphate dehydrogenase (G6PD). In this assay, HK catalyzes the conversion of ATP to ADP and glucose-6-phosphate. Glucose-6-phosphate is then oxidized to D-6-phosphoglucopyranose-1,5-lactone by G6PD, with concomitant reduction of NAD to NADH, which can be monitored at 340 nm. Typical assay solution contained glucose (10 mM), NAD (40 mM), DTT (5 mM), $MgCl_2$ (4 mM), HK (15 unit/ml), G6PD (15 units/ml) and indicated concentrations of ADP and phospho-poly(E_4Y). The reactions were initiated with addition of either VEGFR2Δ50P (600–900 nM) or VEGFR2Δ50 (950 nM). For the phosphorylated enzyme, $K_{M(MgADP)}$ was determined by varying the ADP concentration at fixed concentrations of Mg^{2+} and p-poly(E_4Y) (10 and 8 mM, respectively). For the non-phosphorylated enzyme, $K_{M(MgADP)}$ was determined by varying the ADP concentration at fixed concentrations of Mg^{2+} and p-poly(E_4Y) (10 and 5 mM, respectively). $K_{M(p\text{-}polyE4Y)}$ for the both the phosphorylated and non-phosphorylated enzyme was determined by varying the concentration of p-poly(E_4Y) at fixed concentrations of Mg^{2+} and ADP (10 and 1 mM, respectively).

In vitro *autophosphorylation for crystallization and mass spectrometry*

Aliquots of frozen VEGFR2Δ50 were thawed by immersion in cold H_2O and pooled at 4°C. $MgCl_2$ and ATP were added to 26 mM and 4 mM, respectively. VEGFR2Δ50 was incubated at 4°C for 1 h. This material (VEGFR2Δ50P) was then buffer exchanged into a solution of 10 mM Hepes pH 7.5, 10 mM DTT, and 10 mM NaCl, and was concentrated using a Centriprep-10 (Millipore) to 5 mg protein/ml.

Mass spectrometry: trypsin digestion

Trypsin digestions of purified VEGFR2Δ50 and VEGFR2Δ50P were conducted at 37°C using 0.37 mg/ml protein in 25 mM NH_4HCO_3 at pH 7.7 with a reaction volume of 100 µl for two days.

Mass spectrometry: MALDI MS

Matrix-assisted laser desorption mass spectrometry (MALDI MS) analyses were performed in a Voyager-Elite, time-of-flight mass spectrometer with delayed extraction (PerSeptive Biosystems, Inc., Framingham, MA). A volume of 1 µl of digested protein sample was mixed with 1 µl of matrix (α-cyano-4-hydroxy-cinnamic acid) in a solution of 50% (v/v) acetonitrile and 0.25% (w/w) trifluoroacetic acid in water. Samples were irradiated with a nitrogen laser operated at 337 nm.

Mass spectrometry: nanoESI MS

Nanoelectrospray ionization (NanoESI) MS analyses were performed on a triple quadrapole mass spectrometer (PE Sciex API III, Alberta, Canada) modified with a nanoESI source from Protana A/S (Denmark). The ESI volatage was set at 700 V and the orifice settings were maintained at 100 V. A volume (3 µl) of digested protein was mixed with 7 µl of methanol and 0.5 µl formic acid, and then 4 µl of this sample was injected into the mass spectrometer. Ion scans were used to obtain the sequence of phosphopeptides.

Crystallization

Purified phosphorylated VEGFR2Δ50 was concentrated on average to 5 mg protein/ml using a Centricon-10 centrifugal concentrator. Crystals were grown by the hanging-drop vapor-diffusion method at 4°C. Drops containing 2 µl of protein solution and 2 µl of a mother-liquor solution (100 mM Hepes pH 7.2, 2 M $(NH_4)_2SO_4$, and 2% (v/v) monomethylether polyethylene glycol, molecular weight (MW) ≈ 550 kDa) were equilibrated above a 1 ml reservoir of the mother-liquor solution to which 50 mM β-mercaptoethanol had been added. Crystals appeared after 3–4 days and grew to as large as 0.3 × 0.2 × 0.5 mm over 21 days.

Western-blot analysis of crystals

Crystals were removed from the crystallization drops with fiber loops, washed four times in 5 µl of mother-liquor solution, and dissolved in 25 µl of a solution containing 50 mM Hepes pH 7.5 and 10% DMSO. The crystallization sample was run on an SDS polyacrylamide gel along with control samples of approximately 0.15 µg of purified nonphosphorylated VEGFR2Δ50 and VEGFR2Δ50 after 20 min and 40 min of the autophosphorylation reaction. Kinase activity in all samples was stopped by bringing the final solution to 25 mM Hepes pH 8.0, 40 mM EDTA pH 8.0, 10% glycerol, and 40 mM NaCl. The western blot was performed using the ECL blotting kit from Amersham Inc. (Arlington Heights, IL). The primary antibody solution was a 1:1000 dilution of mouse anti-phosphotyrosine mAb PY20 and the secondary antibody was a 1:1500 dilution of goat anti-mouse horseradish peroxidase conjugate. Both antibodies were obtained from Transduction Laboratories Inc. (Lexington, KY).

Data collection

X-ray diffraction data sets were collected using a Rigaku RU-200 rotating anode X-ray generator (CuKα) operated at 50 kV and 100 mA and equipped with Supper focusing mirrors and a MAR345 MAR Research image plate detector. Data collection on frozen crystals was performed by transferring a crystal into a cryoprotectant solution (100 mM Hepes at pH 7.2, 2.2 M $(NH_4)_2SO_4$, 0.6 M sucrose, 0.55 M glucose, and 2% (v/v) monomethylether polyethylene glycol MW = 550 kDa), flash freezing the crystal in liquid nitrogen, and then transferring the frozen crystal into a stream of nitrogen at −186°C. Data were integrated and scaled using the programs DENZO and SCALEPACK [48] Data-collection statistics are given in Table 2.

Structure determination and refinement

Initial protein phases were obtained using the AMoRe molecular-replacement program [49], molecule 1 of the FGFR1 structure [30] (PDB entry 1FGK) as a search probe, and the native1 data set. The correct solution was achieved by including the FGFR1 sidechains and removing mobile residues of the activation loop (640–660), the N terminus (464–467), a short loop (517–520), and the C termius (760–762) from the search model. The correct solution was the top peak in the rotation and translation functions and had a correlation coefficient of 0.31. Rigid-body refinement in AMoRe improved the solution to a correlation coefficient of 0.49 and an R factor of 46.3% in the 12.0–4.0 Å resolution range. The correctness of this solution was cross-checked by calculation of a difference Fourier with a $KAu(CN)_2$ derivative. This derivative was generated by soaking a crystal for three days in reservoir solution containing 0.5 mM $KAu(CN)_2$ and then increasing the heavy-atom concentration to 5 mM and soaking for an additional 64 h. Scaling of data sets, Patterson calculations, Fourier calculations, and the generation of phases were performed using the program package Xtalview [50].

Refinement of the model was carried out using X-PLOR version 3.1 [51]. Calculation of electron-density maps and model fitting was performed using XtalView [50]. Refinement was begun using a data set collected at 4°C (native2) and was completed using a data set (native3) collected at −186 °C. The final R factor is 20.2% for data in the range 8–2.4 Å ($F_o > 2\sigma$). The average B value for all atoms is 31.8 Å^2 for protein atoms and 42.8 Å^2 for water molecules. The final model includes residues 820–939, 998–1047, and 1064–1168. Of these residues, the sidechains of K838, R842, F845, Y938, K939, D998, K1023, R1027, Y1038, K1039, K1110, and E1113 could not be modeled beyond Cβ because of a lack of interpretable density. Analysis of mainchain torsion angles using PROCHECK [52] shows that, of the 275 residues in the model, none occurs in the disallowed region and only four occur in the generously allowed region of a Ramachandran plot. Water molecules, of which there were 182, were fit to electron-density peaks that were greater than 3σ and were located in positions for making reasonable hydrogen bonds to the protein or other water molecules. Superpositions of various kinase structures were performed using the graphics program Insight II (Molecular Simulations Inc., San Diego, CA) [53].

Accession numbers

The coordinates for VEGFR2Δ50P have been deposited in the Protein Data Bank with accession code 1vr2 and will be available one year after the publication date of this article.

Acknowledgements

We thank the following individuals: Dr S Worland, Dr D Matthews, and L Pelletier for valuable discussions; F Maldonado for help in generating viral stocks; Drs G Siuzdak and J Wu of the Scripps Research Institute (La Jolla, CA) for performing mass spectrometry experiments; J Meador for generating Figure 1; and Dr J Davies for critical reading of the manuscript.

References

1. Folkman, J. & Shing, Y. (1992). Angiogenesis. *J. Biol. Chem.* **267**, 10931-10934.
2. Risau, W. (1995). Differentiation of endothelium. *FASEB J.* **9**, 926-933.
3. Pepper, M.S. (1996). Positive and negative regulation of angiogenesis: from cell biology to the clinic. *Vasc. Med.* **1**, 259-266.
4. Kuiper, R.A.J., Schellens, J.H.M., Blijham, G.H., Beijnen, J.H. & Voest, E.E. (1998). Clinical research on antiangiogenic therapy. *Pharmacol. Res.* **37**, 1–16.
5. Kumar, R. & Fidler, I.J. (1998). Angiogenic molecules and cancer metastasis. *In Vivo* **18**, 27-34.
6. Szekanecz, Z., Szegedi, G. & Koch, A.E. (1998). Angiogenesis in rheumatoid arthritis: pathogenic and clinical significance. *J. Investig. Med.* **46**, 27-41.
7. Tolentino, M.J. & Adamis, A.P. (1988). Angiogenic factors in the development of diabetic iris neovascularization and retinopathy. *Int. Ophthalmol. Clin.* **38**, 77-94.
8. Dvorak, H.F., Brown, L.F., Detmar, M. & Dvorak, A.M. (1995). Vascular permeability factor/vascular endothelial growth factor, microvascular hyperpermeability, and angiogenesis. *Am. J. Pathol.* **146**, 1029–1039.
9. Thomas, K.A. (1996). Vascular endothelial growth factor, a potent and selective angiogenic agent. *Cell* **271**, 603-606.
10. Ferrara, N. & Davis-Smyth, T. (1997). The biology of vascular endothelial growth factor. *Endocrine Rev.* **18**, 4-25.
11. Borgström, P., Hillan, K.J., Sriramarao, P. & Ferrara, N. (1996). Complete inhibition of angiogenesis and growth of microtumors by anti-vascular endothelial growth factor neutralizing antibodies. Novel concepts of angiostatic therapy from intravital videomicroscopy. *Cancer Res.* **56**, 4032-4039.
12. Adamis, A.P., *et al.*, & Miller, J.W. (1996). Inhibition of VEGF prevents retinal ischemia-associated iris neovascularization in a primate. *Arch. Ophthalmol.* **114**, 66-71.
13. Jeltsch, M., *et al.*, & Alitalo, K. (1997). Hyperplasia of lymphatic vessels in VEGF-C transgenic mice. *Science* **276**, 1423–1425.
14. Waltenberger, J., Claesson-Welsh, L., Siegbahn, A., Shibuya, M. & Heldin, C.-H. (1994). Different signal transduction properties of KDR and Flt1, two receptors for vascular endothelial growth factor. *J. Biol. Chem.* **269**, 26988-26995.
15. Seetharam, L., Gotoh, N., Maru, Y., Neufeld, G., Yamaguchi, S. & Shibuya, M. (1995). A unique signal transduction pathway for the FLT tyrosine kinase, a receptor for vascular endothelial growth factor. *Oncogene* **10**, 135–147.
16. Shalaby, F., Rossant, J., Gertsenstein, M., Wu, X.-F., Breitman, M.L. & Schuh, A.C. (1995). Failure of blood island formation and vasculogenesis in Flk-1 deficient mice. *Nature* **376**, 576-579.
17. van der Greer, P., Hunter, T. & Lindberg, R.A. (1994). Receptor protein-tyrosine kinases and their signal transduction pathways. *Annu. Rev. Cell Biol.* **10**, 251-337.
18. Taylor, G.R., Reedijk, M., Rothwell, V., Rohrschneider, L. & Pawson, T. (1989). The unique insert of cellular and viral fms protein tyrosine kinase domains is dispensable for enzymatic and transforming activities. *EMBO J.* **8**, 2029-2037.
19. Heidaran, *et al.*, & Aaronson, S.A. (1991). Deletion or substitution within the a platelet-derived growth factor receptor kinase insert domain: Effects on functional coupling with intracellular signalling pathways. *Mol. Cell. Biol.* **11**, 134–142.
20. Yu, J.C., *et al.*, & Aaronson, S. A. (1991) Tyrosine mutations within the a platelet-derived growth factor receptor kinase insert domain abrogate receptor-associated phosphatidylinositol-3 kinase activity without affecting mitogenic or chemotactic signal transduction. *Mol. Cell. Biol.* **11**, 3780-3785.
21. Kazlauskas, A. & Cooper, J.A. (1989). Autophosphorylation of the PDGF receptor in the kinase insert region regulates interactions with cell proteins. *Cell* **58**, 1121–1133.

22. Lev, S., Givol, D. & Yarden, Y. (1992). Interkinase domain of kit contains the binding site for phosphatidylinositol 3' kinase. *Proc. Natl Acad. Sci. USA* **89**, 678-682.
23. Reedijk, M., *et al.*, & Pawson, T. (1992) Tyr721 regulates specific binding of the CSF-1 receptor kinase insert to PI 3'-kinase SH2 domains: a model for SH2-mediated receptor-target interactions. *EMBO J.* **11**, 1365–1372.
24. Bazenet, C.E., Gelderloos, J.A. & Kazlauskas, A. (1996). Phosphorylation of Tyrosine 720 in the platelet-derived growth factor a receptor is required for binding of Grb2 and SHP-2 but not for activation of Ras or cell proliferation. *Mol. Cell. Biol.* **16**, 6926-6936.
25. Dougher-Vermazen, M., Hulmes, J.D., Bödhlen, P. & Terman, B.I. (1994). Biologicalical Activity and phosphorylation sites of the bacterially expressed cytosolic domain of the KDR VEGF-Receptor. *Biochem. Biophys. Res. Commun.* **205**, 728-738.
26. Knighton, D.R., *et al.*, & Sowadski, J.M. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 407-413.
27. Johnson, L.N., Noble, M.E.M. & Owen, D.J. (1996). Active and inactive protein kinases: structural basis for regulation. *Cell* **85**, 149-158.
28. Hubbard, S.R., Wei, L., Ellis, L. & Hendrickson, W.A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**, 746-754
29. Hubbard, S.R. (1997). Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *Embo J.* **16**, 5572-5581.
30. Mohammadi, M., Schlessinger, J. & Hubbard, S.R. (1996). Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism. *Cell* **86**, 577-87.
31. Mohammadi, M., *et al.*, & Schlessinger, J. (1997). Structures of the tyrosine kinase domain of the fibroblast growth factor receptor in complex with inhibitors. *Science* **276**, 955-960.
32. Parast, C.V., Mroczkowski, B., Pinko, C., Misialek, S., Khambatta, G. & Appelt, K. (1998). Characterization and kinetic mechanism of catalytic domain of human vascular endothelial growth factor receptor-2 tyrosine kinase (VEGFR2 TK), a key enzyme in angiogenesis. *Biochemistry* **37**, 16788–16801.
33. Cox, S., Radzio-Andzelm, E. & Taylor, S.S. (1994). Domain movements in protein kinases. *Curr. Opin. Struct. Biol.* **4**, 893-901.
34. Yamaguchi, H. & Hendrickson., W.A. (1996). Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature* **384**, 484-489.
35. Zheng, J., *et al.*, & Sowadski, J.M. (1993). Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and a peptide inhibitor. *Biochemistry* **32**, 2154–2161.
36. Severinsson, L., Ek, B., Mellström, K., Claesson-Welsh, L. & Heldin, C.-H. (1990). Deletion of the kinase insert sequence of the platelet-derived growth factor β-receptor affects receptor kinase activity and signal transduction. *Mol. Cell. Biol.* **10**, 801-809.
37. Shibuya, M., *et al.*, & Sato, M. (1990) Nucleotide sequence and expression of a novel human receptor-type tyrosine kinase gene (flt) closely related to the fms family. *Oncogene* **5**, 519-524.
38. Gronwald, R.G., *et al.*, & Murray, M.J. (1988). Cloning and expression of a cDNA coding for the human platelet-derived growth factor receptor: evidence for more than one receptor class. *Proc. Natl Acad. Sci. USA* **85**, 3435-3439.
39. Matsui, T., *et al.*, & Aaronson, S. (1989). Isolation of a novel receptor cDNA establishes the existence of two PDGF receptor genes. *Science* **243**, 800-804.
40. Yarden, Y., *et al.*, &, Ulrich, A. (1987). *EMBO J.* **6**, 3341-3351.
41. Kazlauskas, A. & Cooper, J.A. (1989). Autophosphorylation of the PDGF receptor in the kinase insert region regulates interactions with cell proteins. *Cell* **58**, 1121–1133.
42. Terman, B.I., *et al.*, & Bohlen, P. (1992). Identification of the KDR tyrosine kinase as a receptor for vascular endothelial cell growth factor. *Biochem Biophys. Res. Commun.* **187**, 1579-1586.
43. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H.(1992). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. (1986). *Biotechnology* **24**, 17-27.
44. Wei, L., Hubbard, S.R., Hendrickson, W.A. & Ellis, L. (1995). Expression, characterization, and crystallization of the catalytic core of the human insulin receptor protein-tyrosine kinase domain. *J. Biol. Chem.* **270**, 8122-8130.
45. Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA*, **82**, 488-492.
46. Schoner, B.E., Belagaje, R.M. & Schoner, R.G. (1986). Translation of a synthetic two-cistron mRNA in Escherichia coli. *Proc. Natl Acad. Sci. USA*, **83**, 8506-8510.
47. Kan, C., Gehring, M.R., Nodes, B.R., Janson C.A., Almassy, R.J. & Hostomska, Z. (1992). Heterologous expression and purification of active human phosphoribosylglycinamide formyltransferase as a single domain. *J. Protein Chem.* **11**, 467-73.
48. Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing.* (Sawyer L., Isaacs N. & Bailey, S., eds), pp. 56-62, SERC Daresbury Laboratory, England.
49. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157-163.
50. McRee, D.E. (1992). A visual protein crystallographic software system for X11/Xview. *J. Mol. Graph.* **10**, 44-46.
51. Brünger, A.T. (1992). X-PLOR Version 3.1. *A system for X-ray Crystallography and NMR.* Yale University Press, New Haven, CT.
52. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291.
53. INSIGHT II User Guide. (1997). Molecular Simulations Inc., San Diego, CA.